

AD-A036 666

DESMATICS INC STATE COLLEGE PA

F/G 6/11

A STATISTICAL STUDY OF GRAB SAMPLING IN A CLOSED ENVIRONMENT.(U)

FEB 77 D E SMITH, T L KING

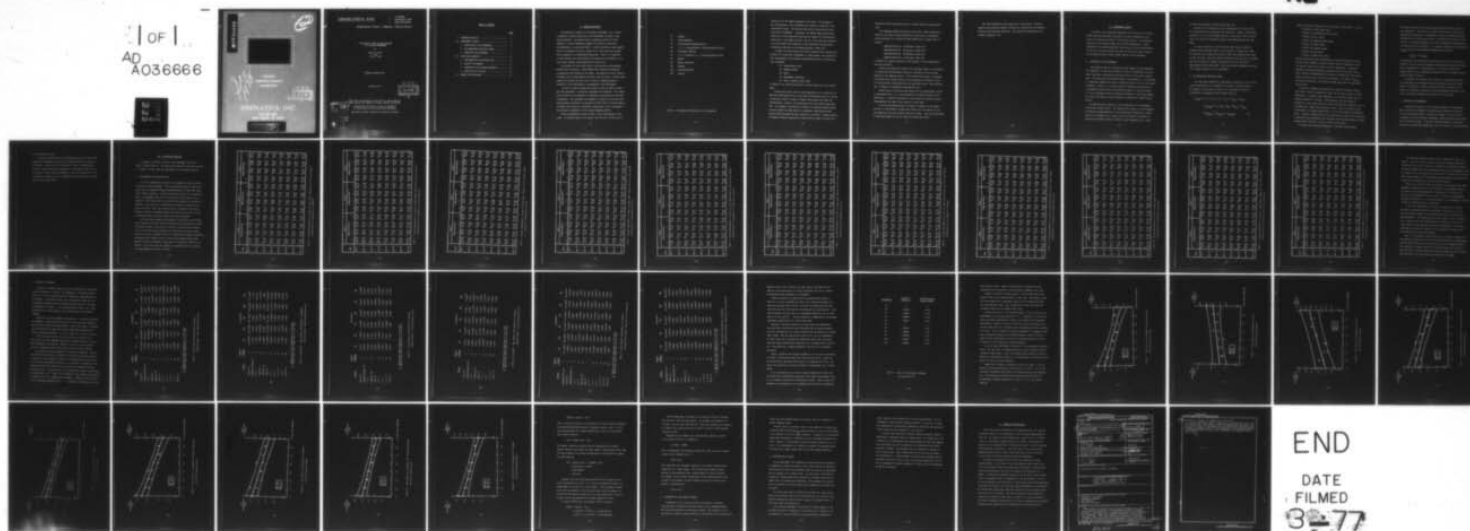
N00014-75-C-1054

UNCLASSIFIED

TR-106-4

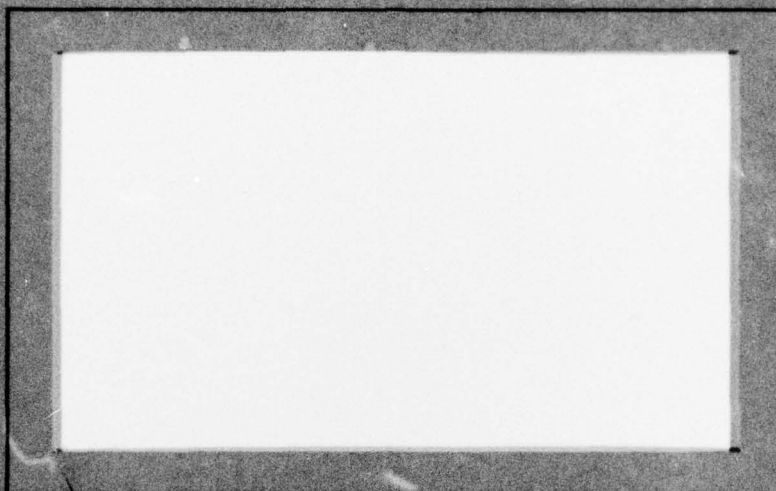
NL

1 of 1  
AD  
A036666



ADA 036666

98  
12



*See  
H 173*

— STATISTICS —  
— OPERATIONS RESEARCH —  
— MATHEMATICS —

DDC  
RECEIVED  
MAR 10 1977  
D

DESMATICS, INC.

P.O. Box 863  
State College, Pa. 16801

DISTRIBUTION STATEMENT A

Approved for public release  
Distribution Unlimited

# DESMATICS, INC.

P. O. Box 863  
State College, Pa. 16801  
Phone: (814) 238-9621

*Applied Research in Statistics - Mathematics - Operations Research*

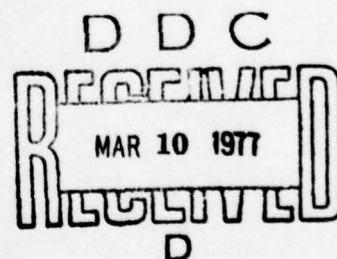
## A STATISTICAL STUDY OF GRAB SAMPLING IN A CLOSED ENVIRONMENT

by

Dennis E. Smith  
and  
Terry L. King

TECHNICAL REPORT 106-4

February 1977



This study was supported by the Office of Naval Research  
under Contract No. N00014-75-C-1054, Task No. NR 042-334

ABSTRACT for	
RTG	White Section <input checked="" type="checkbox"/>
ADD	Buff Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
IDENTIFICATION	
REPRODUCTION/AVAILABILITY CODES	
SPECIAL INFO/SPECIAL	
A	

Reproduction in whole or in part is permitted  
for any purpose of the United States Government

Approved for public release; distribution unlimited



## TABLE OF CONTENTS

	<u>Page</u>
I. PROBLEM DISCUSSION .....	1
II. EXPERIMENTAL DESIGN .....	6
A. CONSTRAINTS ON THE EXPERIMENT .....	6
B. THE UNDERLYING STATISTICAL MODEL .....	7
C. CONDUCTING THE EXPERIMENT .....	9
III. STATISTICAL ANALYSIS .....	11
A. CONSIDERATION OF SUSPICIOUS DATA .....	11
B. ANALYSIS OF VARIANCE .....	24
C. VARIABILITY OF THE SAMPLING METHODS .....	44
D. VARIATION DUE TO FLASKS .....	45
IV. SUMMARY AND CONCLUSIONS .....	47



## I. PROBLEM DISCUSSION

Maintaining the quality of an enclosed environment (e.g., aboard a submarine) requires monitoring of the environment by means of sampling procedures. One general type of sampling procedure is "grab sampling", in which atmospheric samples are obtained, essentially instantaneously, in evacuated flasks. In many situations, these samples may not be analyzed for days or months (e.g., until they are returned from a submarine to a shore-based laboratory). Thus, it is important to know whether such a delay affects the results of the analysis, or in other words, whether these samples have a shelf life.

For purposes of the current study, four particular grab sampling methods were of interest. These methods were defined by differences in preparation and filling of the flasks. One question of major interest is whether any of these methods are better than the others. Another major question is whether the shelf life or choice of sampling method is affected by the concentration of contaminants in the environment.

In order to obtain information on which to base an answer to these and other questions, a statistical experiment was conducted. This experiment consisted of the sampling of atmospheres containing eleven selected contaminants. The experiment was designed to compare the four grab sampling methods of interest, to evaluate the shelf life of the grab samples, to measure the effects of contaminant concentration, and to investigate the sources of variability in the overall sampling process.

Eleven contaminants, listed in Table 1, were investigated in this study. To conserve space in the report, they will for the most part be

#1	Methane
#2	Carbon Monoxide
#3	Dichlorodifluoromethane (R-12)
#4	1, 1, 2, 2-tetrafluoro-1, 2-dichloroethene (R-114)
#5	Vinylidene Chloride
#6	1, 1, 2-trichloro-1, 2, 2-trifluoroethane (R-113)
#7	Hexane
#8	Methyl Chloroform
#9	Benzene
#10	Trichloroethylene
#11	Toluene

Table 1: Contaminants Investigated in the Experiment

referred to by the numbers assigned in the table. For purposes of this investigation, each contaminant was studied at a high and a low concentration level. The specific high and low concentration levels varied with contaminant. In general, the nominal high concentration levels were in the range 60-110 ppm and the nominal low concentration levels were in the range 5-12 ppm. The use of more than one concentration level permits the estimation of any differential effects which concentration may have on the sampling method or shelf life.

A number of sources affect the results obtained from a quantitative analysis of a particular contaminant in a grab sample. For purposes of this investigation, the following contributing sources or factors have been considered:

- (1) Concentration level
- (2) Sampling method
- (3) Flask
- (4) Measurement (analysis)
- (5) Time to analysis (delay time)

In addition, the interaction effects of these sources were also investigated.

Concentration level, which is self-explanatory, is in itself not an important experimental factor, since high concentrations would, of course, be expected a priori to result in higher observations than would low concentrations. However, its interactions with other factors are important. For example, one question is whether one sampling method provides the best results for high levels of contaminant concentrations while another method provides the best results for low levels. Another question is whether different concentration levels have any effect on shelf life.



This would involve interactive effects of delay time with concentration level.

Four sampling methods were used in this study. These methods were defined by the type of flask preparation (vacuum only or vacuum-bake) and the procedure for filling the flask (single fill or purge fill) as follows:

Sampling Method #1: Vacuum-bake, Single fill

Sampling Method #2: Vacuum only, Single fill

Sampling Method #3: Vacuum-bake, Purge fill

Sampling Method #4: Vacuum only, Purge fill

In addition to samples obtained by these methods, in-situ measurements were also obtained.

Because each sample was obtained in a different flask, it is realistic to assume that each individual flask will contribute to the variation observed in the sampling results. This variation may result, for example, from differential effects which occur in preparing or filling the flask or from particular characteristics of an individual flask. This variation may, in general, be regarded as experimental error.

Another source of variation that enters into the experimental results is measurement error, i.e., the differences in repeated analyses of the same flask. To provide an estimate of this variation, duplicate analyses (measurements) were made on the contents of each flask.

To study any possible effects that delay in analysis of the sample may have, it was decided to study these effects over the period from the time sampling occurs until seventy days have elapsed. This was accomplished by analyzing samples at ten day intervals during this period.

The study described in this report was a joint effort. The grab sampling and associated chemical analyses were conducted by the Chemistry Division, Naval Research Laboratory. The statistical analysis was performed by Desmatics, Inc.

## II. EXPERIMENTAL DESIGN

The object of the experiment described in this report was to obtain, within existing operational constraints, the most information on the sources of variation discussed in the previous section. Any experiment, when set up and analyzed, makes use of certain assumptions. A major assumption is the underlying model, which affects not only specification of the data collection process, but also the final statistical analysis. The following sections discuss salient details of the experiment.

### A. CONSTRAINTS ON THE EXPERIMENT

The apparatus used for obtaining the grab samples in this experiment was a manifold. Flasks were attached to ports on the manifold, which was then filled with a specified concentration level of the eleven contaminants. To prevent any unanticipated biases in an experiment, it is desirable to completely randomize the order of sampling relative to all of the factors considered. However, it is easily seen that requiring concentrations to be randomized would involve an extraordinary amount of time and expense. It is more reasonable from an operational standpoint to obtain all samples for a given concentration before samples are obtained for another concentration.

By imposing this constraint on the randomization in the experiment, a split-plot design results. This design owes its name to the development of statistical experimentation in the agricultural fields where a single level of a treatment (e.g., a given type of fertilizer) is applied to a large plot of ground, while all levels of other treatments (e.g., types



of seeds) were applied at random within that plot.

Because of the constraints on the randomization, the actual effect of concentration is not estimated very accurately. However, as mentioned previously, it is not the singular effects of this factor that are important, but rather the interactive effects with sampling method and time delay.

To study the effects of time delay and still keep the number of samples within reasonable limits, it was decided to analyze each flask at intervals over a seventy day period, rather than analyzing each flask at only one time. Thus, the results obtained for any time are not independent of the results at another time because they are both based on measurements from the same flask. This adds another randomization restriction, repeated measures, on the experiment.

#### B. THE UNDERLYING STATISTICAL MODEL

The usual model assumed in an experimental situation is one in which a given observation  $y$  is a composite of a number of additive effects. For example, in the current situation, the usual model (taking the constraints on randomization into account) would be:

$$\begin{aligned}
 y_{ijkmp} = & M + C_i + P_{j(i)} + S_k + CS_{ik} + PS_{jk(i)} + F_{m(ijk)} \\
 & + A_{n(ijkm)} + T_p + CT_{ip} + ST_{kp} + CST_{ikp} + PT_{jp(i)} \\
 & + PST_{jkp(i)} + FT_{mp(ijk)} + \epsilon_{np(ijkm)} \quad (1)
 \end{aligned}$$

where the capital letters refer to the sources of the effects. That is,

M denotes an overall mean effect

C denotes the concentration level effect

P denotes the plot effect

S denotes the sampling method effect

F denotes the flask effect

A denotes the analysis effect

T denotes the time effect

$\epsilon$  denotes a random error effect.

The two- and three-letter combinations in the model indicate the interactive effects between the various factors, while the subscripts denote the specific level of any factor or factor combination. A subscript expression containing parentheses indicates that the particular factor levels outside the parentheses are nested in a hierarchical relationship within the factor levels inside the parentheses. For example,  $PT_{jp(i)}$  indicates the nesting of the  $j$ th plot and  $p$ th time within the  $i$ th concentration level.

An implicit assumption underlying the statistical analysis based on this model is that the random error term  $\epsilon$  has the same variance for every observation. In general, this assumption is not critical unless there is extreme inequality of variance. Unfortunately, this point might be approached in the current situation, since variation may be expected to be on a proportional, rather than an absolute, basis. That is, variation would be expected to be roughly proportional to the true contaminant concentration level. Therefore, observations for low concentration levels would be expected to have much less variability in an absolute sense than observations for high concentrations, and this, in fact, is what the data shows.

To prevent this extreme heterogeneity of variance from adversely

affecting the results of any resulting statistical analysis, there are two alternatives. The first is to analyze the results for high and low concentrations separately. This approach, however, precludes obtaining information on the interactive effects of concentration levels with other factors. The second alternative is to use an appropriate transformation to stabilize the error variance. If, instead of analyzing the observations  $y_{ijkmp}$  directly, the transformed logarithmic observations

$$w_{ijkmp} = \ln y_{ijkmp}$$

are analyzed, the resulting error variance structure will be relatively stable so that the overall experiment involving both concentration levels may be considered in a single analysis. It should be noted that an additive model in the transformed observations (the w's) implies a multiplicative model in the y's.

In general, the existing error structure of the problem indicates that analyzing the transformed observations should provide the optimal amount of information. However, because there may be interest in the results of an analysis conducted on the untransformed observations, pertinent information regarding this analysis is also included.

#### C. CONDUCTING THE EXPERIMENT

As mentioned previously, the constraints imposed on the experiment made a split-plot arrangement involving repeated measures and nesting the most appropriate design. Although this design is not one found in a textbook, it may be obtained by imbedding repeated measures with nesting into a split-plot design. To provide reasonable degrees of freedom for testing for statistically significant sources of variation, three replicates



of this design were used.

To insure an analysis free of unintentional bias, the actual grab sampling and subsequent analysis were randomized within the existing constraints. Specifically, the position of flasks (corresponding to the given sampling types) was randomized on the manifold and the order of analysis at each time was randomized. This was accomplished by the experimenter following a set of randomization plans which were obtained from a random number table.

### III. STATISTICAL ANALYSIS

A summary of the data collected in the experiment is given in Table 2 through Table 12. The statistical analysis of this data consists of a number of steps, which are described in the following sections.

#### A. CONSIDERATION OF SUSPICIOUS DATA

As in most experimental studies of this magnitude, there was suspicious data in this experiment. This suspiciousness existed because this particular data appeared extremely unrepresentative of other data taken under similar conditions. Actually, there were three sets of suspicious data. An assignable cause for the first set was easily found to be a slippage of the decimal point in the data cards supplied for statistical analysis. Thus, this set of suspicious data, which was erroneous, was easily corrected. Only this corrected data is included in the tables. Because no specific assignable causes were found for the other two sets of suspicious data, the tables contain the data as recorded.

A situation more complex than slippage of the decimal point existed in the second set of data, which corresponded to low concentration levels. The suspiciousness of this data was traced to the quantitative analyses of one flask which yielded extremely low observations for most contaminants and, in fact, readings of zero for contaminant #8 (methyl chloroform). The second flask, however, did not confirm these extreme observations. Because of this, an abnormally large amount of variability exists in the results. This can be seen, for example, by examining the low concentration measurements in plot #1 of Table 9.

Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	73.71 8.59	73.43 8.60	73.79 8.68	75.07 8.61	71.63 8.87	72.32 9.52	72.81 8.64	72.08 8.51	70.67 9.22	70.44 9.07	70.90 9.45	70.71 9.71
Day 10	70.93 8.65	71.52 8.64	70.91 8.67	71.34 9.12	71.36 8.93	72.29 8.68	71.88 8.74	72.77 8.77	70.20 8.68	69.81 8.86	69.78 8.61	69.85 8.95
Day 20	72.12 8.78	71.71 8.90	71.90 9.10	71.64 8.86	71.31 9.75	72.58 9.58	72.27 8.69	72.11 8.62	69.03 8.91	70.06 8.81	68.81 9.07	70.23 8.99
Day 30	72.80 8.76	73.02 8.54	72.63 8.87	72.07 8.67	72.72 8.90	72.68 8.68	72.53 8.57	72.31 8.82	70.07 8.91	70.58 8.67	70.61 8.91	71.47 8.89
Day 40	71.97 8.83	72.16 8.73	72.03 8.70	72.92 8.84	71.46 9.04	71.59 9.11	71.76 9.15	71.40 8.76	69.28 8.72	70.16 8.41	71.25 8.63	69.58 8.83
Day 50	70.90 8.95	71.42 8.95	71.84 9.13	71.47 9.16	72.54 8.75	71.72 8.82	71.84 8.78	72.06 8.94	70.31 8.78	71.40 8.81	70.66 9.00	70.87 8.70
Day 60	72.67 9.08	73.03 8.72	73.34 8.68	72.88 9.13	70.62 8.82	70.21 9.33	70.98 9.46	70.30 8.89	70.89 8.56	73.51 8.38	71.94 8.61	72.33 8.54
Day 70	71.56 8.64	72.35 8.88	72.36 9.45	71.76 8.96	70.63 8.66	70.49 8.36	71.60 8.59	71.02 8.89	70.38 8.25	70.23 8.10	70.47 8.39	70.57 8.23

Table 2: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #1 (Methane)



Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	79.34 10.80	79.25 10.56	79.69 10.50	79.42 10.43	74.28 11.23	73.84 12.05	74.11 10.88	74.31 10.64	98.78 11.37	78.02 11.26	67.42 11.56	67.63 12.31
Day 10	78.15 10.82	78.20 10.81	76.57 11.12	79.36 11.22	72.82 10.27	73.80 10.16	72.28 10.22	71.45 10.19	98.55 9.80	76.18 9.79	65.24 10.18	67.46 9.79
Day 20	75.71 10.42	75.12 10.56	75.42 10.66	76.11 10.14	70.42 11.35	70.91 11.30	70.70 10.98	71.37 10.37	95.32 9.71	74.36 9.55	64.29 9.83	66.88 9.73
Day 30	74.78 10.15	74.96 10.11	74.98 10.53	74.59 10.04	70.30 9.87	70.61 9.85	70.06 9.82	70.08 9.94	94.33 9.38	73.32 9.37	63.22 9.54	68.00 9.55
Day 40	73.21 9.65	72.83 9.90	72.63 10.33	72.79 9.80	68.60 9.59	68.31 9.68	68.38 9.81	67.42 9.78	93.62 9.38	72.28 9.41	66.24 9.32	64.67 9.37
Day 50	71.55 9.92	72.33 9.72	71.48 10.17	70.96 9.74	68.04 9.65	67.43 9.56	66.91 9.53	66.26 9.23	96.97 9.35	78.01 9.13	67.14 9.25	68.63 9.06
Day 60	70.65 10.31	70.82 11.09	70.71 11.46	71.13 10.04	66.07 9.21	65.62 9.53	66.47 9.28	66.71 9.24	100.09 9.01	80.55 8.98	68.72 9.05	70.51 9.10
Day 70	69.01 9.20	67.66 9.40	69.10 11.62	68.67 9.45	65.42 9.22	64.52 9.08	64.98 9.12	66.02 9.16	99.62 8.99	78.38 8.73	68.66 9.07	69.33 9.02

Table 3: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #2 (Carbon Monoxide)

Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	83.94 10.56	83.24 10.83	83.32 11.00	83.99 10.74	81.93 9.65	81.75 9.54	81.63 9.59	81.49 9.71	83.59 9.76	83.66 9.66	83.55 9.71	83.75 9.60
Day 10	79.83 9.60	81.30 9.72	79.67 9.76	78.85 9.88	80.47 9.65	80.51 9.83	80.44 9.77	80.63 9.70	84.61 9.56	84.37 9.49	83.74 9.50	84.54 9.48
Day 20	81.52 9.67	81.22 9.93	81.33 9.64	80.90 10.11	80.46 9.64	80.66 9.62	80.40 9.68	80.39 9.70	82.09 9.67	83.66 9.49	82.08 9.60	84.47 9.68
Day 30	81.62 9.63	81.80 9.54	81.63 9.87	81.64 9.49	81.28 9.33	81.23 9.65	81.20 9.63	81.06 9.47	83.72 9.86	84.18 9.53	84.22 9.70	85.51 9.48
Day 40	81.43 9.69	80.68 10.37	80.73 9.63	81.84 9.54	84.44 9.55	83.98 9.60	84.45 9.55	84.19 9.67	81.72 10.38	82.00 9.93	82.79 10.03	81.55 10.84
Day 50	80.40 9.49	80.68 9.53	80.26 9.47	80.27 9.52	84.65 9.93	84.47 10.41	84.32 9.94	84.35 10.76	85.50 9.95	86.05 9.90	85.83 9.92	86.38 9.88
Day 60	81.00 9.58	81.21 9.59	81.68 9.67	81.31 9.51	83.74 9.88	83.17 10.09	84.20 9.97	83.30 9.96	86.52 9.86	88.09 9.94	86.87 10.01	87.86 9.90
Day 70	84.19 10.08	84.23 10.74	84.50 11.08	83.88 11.69	84.64 9.71	84.28 9.90	85.06 9.86	83.43 9.84	85.71 10.84	85.57 11.48	85.74 10.50	85.47 10.37

Table 4: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #3 (R-12)

Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	111.26 8.62	109.42 8.55	109.91 8.56	111.83 8.56	100.85 8.26	101.28 8.17	100.51 8.16	100.17 8.24	106.33 8.35	106.49 8.37	106.25 8.18	106.46 8.28
Day 10	103.68 8.30	104.56 8.30	102.28 8.44	103.64 8.25	103.11 8.24	103.14 8.28	102.99 8.28	103.05 8.28	108.24 8.05	108.21 8.09	107.16 8.05	108.45 8.02
Day 20	99.88 8.21	99.95 8.29	100.06 8.23	99.51 8.34	98.50 8.29	98.70 8.33	98.12 8.20	98.08 8.28	104.25 8.23	107.14 8.17	104.72 8.24	108.60 8.19
Day 30	100.13 8.28	100.32 8.18	100.38 8.21	100.65 8.22	100.06 8.10	99.83 8.16	99.66 8.08	99.52 8.07	108.44 8.49	108.19 8.11	108.27 8.47	109.59 8.09
Day 40	105.02 8.20	103.79 8.14	103.33 8.22	106.53 8.25	109.55 8.21	106.85 8.22	107.63 8.22	106.72 8.28	103.44 9.05	103.92 8.95	105.07 8.97	103.96 9.05
Day 50	98.29 8.10	98.92 8.08	98.28 8.09	98.12 8.08	112.93 8.87	112.63 8.96	112.49 8.88	112.50 8.84	111.55 8.89	112.31 8.89	112.02 8.93	112.83 8.81
Day 60	99.88 8.13	99.69 8.13	100.50 8.17	100.23 8.08	106.83 8.91	105.96 9.00	107.95 8.96	106.17 8.87	113.82 8.79	115.89 8.85	113.67 8.86	115.63 8.84
Day 70	111.60 8.94	111.86 8.98	111.89 9.05	111.16 9.06	108.79 8.70	107.69 8.72	109.15 8.71	106.69 8.72	111.40 9.04	111.19 9.03	111.40 9.15	110.67 9.18

Table 5: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #4 (R-114)



Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	78.85 6.86	78.62 6.70	78.78 7.36	79.41 6.88	79.73 6.82	79.27 6.73	79.85 6.74	79.71 6.71	76.40 6.76	76.42 6.75	76.43 6.66	76.90 6.81
Day 10	77.59 6.71	79.29 6.71	78.92 7.32	78.67 6.72	78.99 6.80	78.99 6.81	78.65 6.75	79.31 6.79	76.72 6.73	77.17 6.68	76.72 6.79	77.80 6.59
Day 20	79.06 6.78	79.15 6.74	78.85 6.66	79.02 6.72	79.37 6.76	78.93 6.73	79.14 6.70	79.07 6.90	76.30 6.77	77.48 6.71	75.71 6.81	78.68 6.69
Day 30	79.68 6.72	80.07 6.79	79.31 6.15	78.24 6.75	79.46 6.71	79.27 6.70	79.17 6.68	79.24 6.68	76.77 6.68	77.68 6.68	77.00 6.80	79.50 6.66
Day 40	79.20 6.65	79.11 5.11	78.70 5.80	78.42 6.69	76.94 6.73	76.80 6.71	76.92 6.72	77.03 6.70	76.71 6.57	77.54 6.53	77.82 6.65	79.03 6.55
Day 50	78.51 6.66	78.72 6.81	77.85 5.38	77.81 6.66	77.47 6.57	76.96 6.57	77.10 6.52	76.93 6.56	74.62 6.47	75.21 6.54	74.71 6.66	77.77 6.54
Day 60	78.99 6.73	79.82 6.88	79.49 5.17	78.20 6.81	76.26 6.49	75.96 6.46	76.51 6.51	75.69 6.46	74.61 6.52	76.24 6.53	75.09 6.61	78.21 6.45
Day 70	77.28 6.70	77.19 6.69	76.28 4.73	75.56 6.52	76.41 6.48	76.01 6.44	76.66 6.47	75.10 6.49	75.00 6.54	75.65 6.50	74.76 6.66	77.86 6.60

Table 6: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #5 (Vinylidene Chloride)

Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	87.18 6.42	86.35 6.29	86.63 6.20	87.62 6.45	85.75 8.58	85.65 8.41	85.76 8.35	85.42 8.37	85.13 8.59	85.03 8.49	85.13 8.37	85.50 8.50
Day 10	83.76 8.40	85.04 8.33	84.10 8.33	84.25 8.36	84.58 5.76	84.47 5.85	84.23 5.80	84.78 5.81	85.83 6.58	86.04 7.03	85.99 6.41	86.06 6.49
Day 20	85.22 5.81	84.82 5.73	84.88 5.67	85.23 5.84	84.80 8.49	84.59 9.31	84.60 8.46	84.74 8.68	84.21 5.88	85.84 5.79	83.91 5.79	86.59 5.97
Day 30	85.60 5.84	85.74 5.79	85.55 5.83	84.89 5.83	85.43 5.89	85.07 5.83	84.87 5.76	85.71 8.06	85.86 6.05	86.36 6.08	86.10 6.30	87.11 5.90
Day 40	85.55 8.69	84.63 7.51	84.61 7.49	85.61 7.34	85.87 5.79	85.58 5.84	86.32 5.81	85.97 5.95	84.10 5.97	84.82 5.96	85.66 5.89	84.65 5.98
Day 50	84.44 5.73	84.16 5.75	83.96 5.93	84.06 5.74	86.51 7.62	86.02 7.52	86.10 6.65	86.34 9.59	85.62 5.90	85.68 5.88	85.92 5.94	86.59 5.86
Day 60	85.01 7.26	85.27 8.89	85.66 8.42	85.16 6.69	85.22 5.84	84.65 5.96	85.72 5.92	85.14 5.89	86.37 5.88	88.13 5.90	86.73 5.86	87.73 5.86
Day 70	86.18 5.92	85.95 5.94	85.44 7.19	85.69 6.03	85.80 5.81	85.22 5.83	86.04 5.83	84.68 5.83	85.89 6.10	85.86 5.97	85.84 6.11	85.59 6.14

Table 7: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #6 (R-113)

Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	103.01 10.40	102.67 10.18	103.72 9.97	104.09 10.57	103.44 10.46	102.99 10.33	103.71 10.34	103.95 10.43	99.98 10.57	100.34 10.36	99.95 10.28	100.79 10.45
Day 10	100.50 10.33	102.44 10.30	102.24 9.74	102.17 10.32	102.64 10.48	102.74 10.49	102.04 10.37	102.54 10.52	100.25 10.38	101.31 10.27	100.37 10.33	100.90 10.32
Day 20	102.56 10.43	102.56 10.19	102.71 9.60	102.52 10.48	103.49 10.40	102.34 10.41	102.74 10.37	102.05 10.41	99.59 10.38	101.03 10.31	98.39 10.40	100.72 10.24
Day 30	103.09 10.35	103.24 10.18	102.78 9.57	101.67 10.28	102.86 10.30	103.07 10.33	102.60 10.27	102.75 10.37	100.41 10.25	101.30 10.21	100.54 10.49	101.74 10.21
Day 40	102.62 10.31	101.86 10.13	101.89 9.31	101.71 10.23	101.13 10.30	100.79 10.51	100.97 10.28	101.46 10.28	99.83 10.19	100.64 10.20	101.64 10.24	100.27 10.27
Day 50	101.87 10.21	101.48 10.21	100.46 9.28	101.26 10.18	101.70 10.23	100.82 10.21	101.04 10.14	101.42 10.15	98.33 10.02	98.97 10.10	98.19 10.23	99.93 10.18
Day 60	102.58 10.29	102.86 10.31	103.33 9.31	101.74 10.30	99.70 10.11	99.50 10.11	100.43 10.10	99.68 10.05	98.31 10.02	100.50 10.12	98.83 10.05	99.79 9.99
Day 70	101.93 10.08	100.53 10.15	100.13 9.40	99.36 10.24	99.85 10.11	99.38 9.99	100.48 10.04	98.38 10.06	98.80 10.19	99.55 10.10	98.35 10.21	98.92 10.23

Table 8: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #7 (Hexane)



Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	66.65 5.38	66.38 5.07	67.10 4.02	67.32 5.32	66.82 5.10	66.50 5.02	67.10 5.05	67.45 5.02	64.38 5.00	64.92 5.02	64.70 4.97	65.13 5.02
Day 10	65.50 5.00	66.05 4.90	66.22 2.72	66.25 5.05	66.40 5.02	66.25 5.02	65.92 5.07	66.20 5.05	64.72 4.97	65.65 4.85	65.10 4.95	64.40 5.00
Day 20	66.25 4.90	65.90 4.90	66.30 2.55	66.47 5.02	66.92 5.00	65.90 5.00	66.45 5.00	65.70 5.00	65.77 4.95	65.22 4.95	63.52 4.92	63.82 4.92
Day 30	66.60 4.95	66.22 4.80	66.35 2.52	65.95 4.97	66.40 5.00	66.72 5.00	66.35 5.02	66.27 5.02	64.87 4.95	65.40 4.92	64.97 5.05	64.13 4.95
Day 40	66.52 4.95	65.20 4.77	65.77 2.50	66.05 4.95	65.25 4.95	64.92 5.00	65.45 4.97	65.57 4.97	64.40 4.85	64.60 4.90	65.72 4.85	62.50 5.05
Day 50	66.05 4.87	65.02 4.77	65.07 2.50	65.80 4.92	65.67 4.92	64.87 5.00	65.32 4.92	65.72 4.95	63.55 4.82	63.60 4.82	63.45 4.90	62.05 5.00
Day 60	66.42 4.95	65.90 4.85	66.92 2.52	66.25 5.00	64.25 4.85	64.02 4.92	64.92 4.95	64.80 4.82	63.40 4.88	64.95 4.92	63.68 4.75	61.63 4.82
Day 70	66.15 4.88	64.27 4.90	64.80 2.47	64.70 5.10	64.65 4.77	64.00 4.82	65.00 4.85	63.77 4.82	63.75 5.00	63.92 4.97	63.55 4.90	60.72 5.02

Table 9: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #8 (Methyl Chloroform)

Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	87.52 7.25	87.07 7.00	89.25 6.52	89.00 7.50	82.97 7.30	82.30 7.20	83.52 7.27	84.27 7.22	80.32 7.42	80.88 7.52	79.77 7.40	80.92 7.55
Day 10	85.45 7.20	87.17 7.05	87.72 6.35	87.17 7.25	87.95 7.32	87.92 7.27	87.07 7.22	87.77 7.27	80.10 7.17	81.87 7.07	80.22 7.20	80.72 7.17
Day 20	81.97 7.17	82.32 6.90	82.50 6.27	82.47 7.17	83.63 7.42	81.80 7.52	82.92 7.52	81.77 7.55	80.13 7.15	81.15 7.10	78.07 7.15	80.22 7.07
Day 30	82.45 7.12	82.35 6.95	82.20 6.25	81.10 7.15	82.52 7.17	83.07 7.25	82.80 7.17	83.30 7.22	80.15 7.13	81.15 7.12	79.65 7.45	80.90 7.13
Day 40	87.65 7.07	86.57 6.97	87.15 6.17	86.70 7.12	83.35 7.17	80.50 7.12	80.97 7.12	81.88 7.10	79.60 7.02	80.67 7.05	81.17 7.12	80.10 7.00
Day 50	81.82 7.02	80.92 6.92	80.10 6.22	81.45 7.05	86.57 7.05	85.10 7.12	85.75 7.05	87.25 7.05	77.90 6.87	78.80 6.92	77.42 7.10	79.85 7.07
Day 60	82.72 7.20	82.85 7.20	83.70 6.32	82.15 7.22	80.05 6.95	80.07 6.97	81.13 6.95	81.30 6.97	76.95 6.92	79.50 7.02	77.30 6.95	78.65 6.90
Day 70	87.17 6.95	84.90 6.92	84.97 6.22	84.35 7.12	79.52 6.90	78.63 6.85	80.40 6.92	78.88 6.95	77.63 7.07	78.67 7.00	76.95 7.07	77.92 7.17

Table 10: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #9 (Benzene)

Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	86.15 8.70	85.65 8.27	88.32 8.22	88.67 9.05	85.60 8.32	84.13 8.27	86.22 8.22	87.50 8.22	81.22 8.40	83.13 8.32	80.72 8.30	83.02 8.37
Day 10	92.97 8.12	85.88 8.05	86.12 7.57	85.67 8.22	84.85 8.55	84.25 8.47	83.90 8.37	83.95 8.50	82.00 8.17	84.35 8.02	81.17 8.17	82.17 8.07
Day 20	83.77 8.25	83.72 7.90	84.57 7.52	83.00 8.40	86.50 8.20	84.07 8.40	85.70 8.35	83.72 8.35	81.95 8.20	83.25 8.25	78.92 8.25	81.88 8.02
Day 30	84.38 8.20	84.07 8.07	83.70 7.50	82.15 8.22	83.38 8.15	83.97 8.22	84.17 8.15	84.95 8.20	81.75 8.15	83.27 8.07	80.72 8.72	82.45 8.10
Day 40	83.30 8.07	82.30 7.97	82.88 7.27	82.15 8.05	82.50 8.25	82.00 8.25	82.63 8.17	83.92 8.07	80.80 8.15	82.37 8.15	82.15 8.22	81.17 8.15
Day 50	82.80 8.10	81.30 8.00	79.77 7.32	82.80 8.05	87.22 8.05	85.22 8.02	84.70 8.05	87.05 8.02	79.35 7.92	80.82 8.10	78.22 8.25	81.67 8.32
Day 60	83.92 8.20	83.90 8.17	84.75 7.25	82.55 8.07	80.90 8.05	80.35 8.15	82.63 8.17	82.17 8.02	78.85 8.07	82.35 8.37	79.05 8.10	81.15 8.00
Day 70	84.07 7.92	81.22 8.07	81.25 7.45	80.00 8.30	80.77 7.97	79.42 7.95	82.25 8.02	80.45 8.10	79.52 8.12	81.22 8.30	78.70 8.17	79.77 8.10

Table 11: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #10 (Trichloroethylene)



Time	PLOT 1				PLOT 2				PLOT 3			
	Sampling Method				Sampling Method				Sampling Method			
	1	2	3	4	1	2	3	4	1	2	3	4
Day 0	94.45 7.07	93.70 6.17	98.52 4.63	98.20 7.52	96.02 6.82	94.67 6.85	97.42 6.75	100.52 6.87	91.45 6.82	94.85 6.85	90.27 6.75	93.47 7.00
Day 10	90.97 6.60	92.77 6.15	95.05 3.90	94.02 6.80	96.22 7.10	95.07 7.10	94.35 6.97	94.60 7.12	92.45 6.70	96.80 6.52	91.60 6.70	92.42 6.65
Day 20	93.47 6.72	93.22 5.97	95.70 3.80	93.25 6.85	98.72 6.60	94.13 7.02	97.55 6.92	93.12 6.92	93.07 6.65	94.37 6.72	87.75 6.70	90.62 6.60
Day 30	94.17 6.67	93.07 6.02	93.77 3.72	91.80 6.75	93.37 6.65	95.02 6.88	95.47 6.70	96.15 6.80	91.67 6.70	93.65 6.57	89.42 7.37	91.22 6.72
Day 40	93.42 6.52	91.00 6.15	92.97 3.47	91.12 6.67	93.52 6.70	92.25 6.77	94.12 6.70	95.22 6.62	90.35 6.72	92.90 6.67	91.77 6.75	90.30 6.60
Day 50	92.65 6.55	89.97 5.95	89.02 3.45	92.67 6.57	94.12 6.63	91.02 6.70	92.80 6.75	96.07 6.67	88.42 6.42	91.02 6.52	86.45 6.77	91.65 6.90
Day 60	94.35 6.75	93.55 6.32	95.90 3.57	91.92 6.75	90.50 6.55	89.75 6.75	93.47 6.70	93.30 6.70	87.85 6.40	92.62 6.72	87.87 6.47	90.60 6.55
Day 70	96.32 6.45	90.50 6.17	90.63 3.55	88.85 6.80	90.52 6.45	88.00 6.50	92.95 6.62	89.37 6.70	88.62 6.62	91.07 6.47	86.95 6.52	88.95 6.70

Table 12: Average Measurements (in ppm) for High (Upper Figure) and Low (Lower Figure) Concentrations of Contaminant #11 (Toluene)

Two different assumptions may be made to explain this large variability. The first is that the data in question indicates experimental variability, and as such, should be included in the statistical analysis. The second assumption is that something went wrong in the experiment, with the observed data reflecting the spurious effect of whatever did go wrong. Under this assumption, the flask generating the suspicious data should be dropped from consideration in the statistical analysis.

Based on the fact that these low readings from one particular flask immediately caused doubt as to their accuracy, the subsequent analysis was based on the second assumption. Unfortunately, this affects the balance of the experiment so that in order to obtain an uncontaminated statistical analysis (i.e., one in which there is no confounding of effects), a number of observations must be discarded.

In the third suspicious set of data, all observations of carbon monoxide for sampling method #1 (vacuum-bake, single fill) in plot #3 were substantially higher than observations for any of the other three sampling methods. This suspicious data was confined to this contaminant only. However, these high observations were repeated for each of two flasks which had been prepared and analyzed independently. Furthermore, investigation uncovered no anomaly in the experimental procedure which would explain these suspicious readings.

Thus, it was decided to analyze the data as recorded, realizing that the suspicious observations would contribute substantially to experimental variation (specifically to the interaction of plot with sampling method). For comparison, a similar analysis was performed using only the data of plots #1 and #2. Although there were numerical differences in the results, the conclusions from each analysis were essentially the same.

## B. ANALYSIS OF VARIANCE

Analysis of variance breaks out the total variation in the experiment and ascribes it to each of the components in the additive model given by (1), or by its multiplicative counterpart, depending upon the one adopted. As previously noted, each concentration (high and low) was analyzed separately under the additive model because of the large heterogeneity of variance. For the multiplicative model (i.e., using logarithms of the observations), both concentrations were analyzed together.

Tables 13, 14, and 15 summarize the analyses. It will be noted that no estimates of flask affects are given for the two analyses (Tables 14 and 15) which include low concentrations. This is the result of discarding the second set of suspicious data and additional observations to prevent confounded estimates. Of course, Table 15 is of major interest.

As can be seen from Table 15, there is little statistical evidence that any of the sampling methods give different results for any of the contaminants investigated. However, for most of the contaminants there is a significant time effect, which warrants further analysis.

The time effect may be examined in a regression framework. That is, the observed concentration is assumed to vary about a mean value which is a function of time, defined by a point on a regression (time trend) line. In the current experimental situation, it is realistic to assume that any time effect should be reflected in a decay process, hopefully the same for high and low. In other words, it is assumed that observed concentration is proportional to  $e^{\alpha t}$ , where  $t$  is the number of days elapsed from sampling to analysis and  $\alpha$  is an unknown decay parameter. Because of this, the use of the transformed observations (i.e., logarithms) is indicated, since this



Source	Degrees of Freedom	Contaminant					
		#1	#2	#3	#4	#5	#6
Plots (P)	2	103.5323**	1888.8060**	273.0729**	1009.0690**	115.7852**	10.9161**
Sampling Method (S)	3	3.4029*	2195.0250	0.3048	1.4159	6.1046	1.0459
P x S	6	0.5368	2118.0230**	2.3140*	7.3418**	17.4132	2.2438
Flask (F)	12	0.5531	38.4875**	0.6358	1.3686*	9.0515**	1.2120*
Analysis (A)	24	0.3221	1.7600	0.4210	0.5572	0.3576	0.5124
Time (T)	7	10.3763	218.6380	61.9307	369.6584	39.2451*	7.4398
P x T	14	10.3575**	113.8769**	34.9628**	314.7861**	10.4058**	10.0670**
S x T	21	0.4984	2.8010	0.8681	2.2932	0.5167	0.7855
P x S x T	42	1.1347	2.6120	0.8728	2.2953	1.0743	0.8093
F x T	84	0.8633**	2.1187	0.7835**	2.1622**	1.0290**	0.9430**
Error (ε)	168	0.4918	1.8290	0.2938	0.4344	0.3368	0.4233

\* Denotes mean square significant at 0.05 level

\*\* Denotes mean square significant at 0.01 level

Table 13: Mean Squares from Analysis of Variance  
(Additive Model - High Concentration Only)

<u>Source</u>	<u>Degrees of Freedom</u>	<u>Contaminant</u>				
		<u>#7</u>	<u>#8</u>	<u>#9</u>	<u>#10</u>	<u>#11</u>
Plots (P)	2	165.3993**	128.4878**	821.2958**	268.9226**	310.3059**
Sampling Method (S)	3	0.9143	5.4178	1.4650	3.7024	1.7110
P x S	6	5.4408	8.1485	10.8160	25.9926*	82.1750
Flask (F)	12	3.5645**	14.2945**	4.1298*	5.6373	27.6121**
Analysis (A)	24	0.9080	0.4866	1.4670	3.5537	7.7890
Time (T)	7	42.4879**	24.3910**	119.3050	83.3876*	125.8210**
P x T	14	8.0249**	2.8434*	71.8050**	24.4271**	17.7490
S x T	21	1.8722	1.6915*	2.9660	6.2866*	15.4890*
P x S x T	42	1.1958	0.8770	1.8350	3.1323	8.0580
F x T	84	1.5641*	1.3252**	2.2938	4.3206	10.2833
Error (ε)	168	0.9791	0.5492	1.6290	3.3696	8.2770

\* Denotes mean square significant at 0.05 level

\*\* Denotes mean square significant at 0.01 level

Table 13 (continued): Mean Squares from Analysis of Variance

(Additive Model - High Concentration Only)

Source	Degrees of Freedom	Contaminant					
		<u>#1</u>	<u>#2</u>	<u>#3</u>	<u>#4</u>	<u>#5</u>	<u>#6</u>
Plots (P)	2	0.0172	5.6358**	0.7868**	0.6191**	0.0403*	2.5078**
Sampling Method (S)	3	0.1982	0.0783	0.7709	0.0000	0.0219	0.1169
P x S	6	0.1036	0.4001	0.1875	0.0039	0.0302*	0.4568
Analysis (A)	12	0.1415	0.1753	0.0898	0.0042	0.0075	0.1955
Time (T)	7	0.4068	8.2592*	1.2509	1.7602	0.1580*	9.2130
P x T	14	0.3867**	2.0832**	0.9591**	0.7090**	0.0436**	9.5849**
S x T	21	0.1395	0.2422	0.1530	0.0149	0.0075	0.1992
P x S x T	42	0.2057*	0.2457	0.2969**	0.0101**	0.0053	0.1836
Error (E)	84	0.1112	0.1995	0.0692	0.0028	0.0045	0.1726

\* Denotes mean square significant at 0.05 level

\*\* Denotes mean square significant at 0.01 level

Table 14: Mean Squares from Analysis of Variance  
(Additive Model - Low Concentration Only)



<u>Source</u>	<u>Degrees of Freedom</u>	<u>Contaminant</u>				
		<u>#7</u>	<u>#8</u>	<u>#9</u>	<u>#10</u>	<u>#11</u>
Plots (P)	2	0.0714*	0.2454**	0.0565*	0.0663	0.6798**
Sampling Method (S)	3	0.0209	0.0807	0.1222	0.1841	1.0265
P x S	6	0.0442*	0.1140**	0.0578**	0.1525**	0.8260**
Analysis (A)	12	0.0142	0.0057	0.0104	0.0310	0.0535
Time (T)	7	0.3823**	0.1542**	0.4281*	0.5347*	0.5018*
P x T	14	0.0390**	0.0359**	0.1312**	0.1822**	0.1734**
S x T	21	0.0245	0.0107	0.0146	0.0511	0.0674
P x S x T	42	0.0198	0.0074	0.0117	0.0518*	0.0569
Error (ε)	84	0.0162	0.0069	0.0126	0.0327	0.0388

\* Denotes mean square significant at 0.05 level

\*\* Denotes mean square significant at 0.01 level

Table 14 (continued): Mean Squares from Analysis of Variance  
(Additive Model - Low Concentration Only)

Source	Degrees of Freedom	Contaminant					
		#1	#2	#3	#4	#5	#6
Concentration (C)	1	421.7026**	384.4001**	435.7341**	612.2873**	579.1592**	635.2274**
Plots (P)	4	0.0053**	0.0053**	0.0129**	0.0263**	0.0032**	0.0247**
Sampling Method (S)	3	0.0017	0.0768*	0.0036	0.0001	0.0008	0.0015
C x S	3	0.0016	0.0772	0.0034	0.0000	0.0022	0.0013
P x S	12	0.0007	0.0735**	0.0010	0.0002**	0.0023**	0.0044*
Analysis (A)	24	0.0008	0.0008	0.0005	0.0001	0.0001	0.0019
Time (T)	7	0.0036	0.0858**	0.0136*	0.0386*	0.0062**	0.0962
C x T	7	0.0025	0.0117	0.0029	0.0023	0.0002	0.0934
P x T	28	0.0030**	0.0148**	0.0058**	0.0120**	0.0011**	0.0954**
S x T	21	0.0009	0.0011	0.0008	0.0002	0.0001	0.0021
C x S x T	21	0.0008	0.0013	0.0008	0.0002	0.0002	0.0019
P x S x T	84	0.0013**	0.0013	0.0014**	0.0002**	0.0002**	0.0019
Error (E)	168	0.0006	0.0010	0.0004	0.0001	0.0001	0.0017

\* Denotes mean square significant at 0.05 level

\*\* Denotes mean square significant at 0.01 level

Table 15: Mean Squares from Analysis of Variance

(Multiplicative Model - High and Low Concentration)

<u>Source</u>	<u>Degrees of Freedom</u>	<u>Contaminant</u>				
		<u>#7</u>	<u>#8</u>	<u>#9</u>	<u>#10</u>	<u>#11</u>
Concentration (C)	1	502.1978**	636.8233**	574.0371**	514.3004**	665.0485**
Plots (P)	4	0.0036**	0.0153**	0.0265**	0.0062**	0.0111**
Sampling Method (S)	3	0.0000	0.0005	0.0006	0.0009	0.0063
C x S	3	0.0005	0.0079	0.0023	0.0030	0.0204
P x S	12	0.0004**	0.0048**	0.0011**	0.0023**	0.0126**
Analysis (A)	24	0.0001	0.0002	0.0002	0.0004	0.0010
Time (T)	7	0.0055**	0.0091**	0.0140**	0.0151**	0.0217**
C x T	7	0.0004	0.0006	0.0032	0.0006	0.0004
P x T	28	0.0005**	0.0009**	0.0034**	0.0019**	0.0022**
S x T	21	0.0003	0.0004	0.0004	0.0008	0.0018*
C x S x T	21	0.0002	0.0007*	0.0002	0.0006	0.0011
P x S x T	84	0.0002	0.0003*	0.0002	0.0007	0.0010
Error (ε)	168	0.0001	0.0002	0.0002	0.0005	0.0009

\* Denotes mean square significant at 0.05 level

\*\* Denotes mean square significant at 0.01 level

Table 15 (continued): Mean Squares from Analysis of Variance

(Multiplicative Model - High and Low Concentration)



permits a direct test of whether the time trend is the same for both high and low concentrations of a given contaminant and, if so, whether an exponential decay assumption is reasonable.

From the analysis of variance for the transformed data given in Table 15, it can be concluded that there is no interaction between concentration and time (the CT term). From this it follows that any time trend which exists is the same for both high and low concentrations. With this conclusion, the next step is to investigate whether or not the time trend is of the form  $e^{\alpha t}$ . This is equivalent to examining the transformed logarithmic observations for a linear time trend.

Because of the equal spacing on the time scale, the experimental variation which resulted from time differences may be easily decomposed into that accounted for by a linear trend and that accounted for by higher-order trends. This was done with the result that, for all contaminants for which there was a statistically significant time trend, the linear trend was always statistically significant but the higher-order trend was not. This means that a simple exponential time trend is a reasonable assumption.

Figure 2 exhibits the estimated parameter  $\alpha$  for the nine contaminants for which a statistically significant time trend was found. (There was no statistically significant time effect for contaminant #1 or #6.) The figure also indicates the predicted change in concentration over a 70-day period.

It is interesting to note that although contaminants #3 (R-12) and #4 (R-114) have statistically significant time trends, the estimated value of  $\alpha$  is positive, indicating an exponential increase. Both of these contaminants are refrigerants, as is contaminant #6 for which no significant

<u>Contaminant</u>	<u>Estimated Value of <math>\alpha</math></u>	<u>Predicted Change after 70 Days</u>
#1	---	---
#2	-.001614	-10.7%
#3	+.000404	+ 2.9%
#4	+.000896	+ 6.5%
#5	-.000425	- 2.9%
#6	---	---
#7	-.000424	- 2.9%
#8	-.000546	- 3.7%
#9	-.000682	- 4.7%
#10	-.000697	- 4.8%
#11	-.000825	- 5.6%

Figure 2: Value of the Estimated Parameter  
 $\alpha$  in Time Trend  $e^{\alpha t}$

time trend was found. Based on these results, it appears that the refrigerants react differently from the other contaminants over time.

Figures 3 through 11 present graphs of the nine significant time trends as well as the observed values at each time. The observed values here are averages (on a logarithmic scale) over all samples and plots for each particular time. Also included are the upper and lower 95% confidence bounds on the estimated time trend.

Because the value of  $\alpha$  is relatively small,  $e^{\alpha t}$  may be fairly well-approximated by the linear equations  $1 + \alpha t$ . Therefore, the time trend would also appear approximately linear in the original (untransformed) observations. However, in order to take the initial concentration into account, there would be a different slope corresponding to each concentration level. This is avoided by using the exponential. Furthermore, when both concentration levels may be analyzed together, there are twice as many effective observations on which the resulting estimates are based. In general, this provides more accurate estimates.

As can be seen from Figures 3 through 11, the experimental determination of the time trend is reasonably accurate, as the relatively narrow confidence bounds attest. Thus, the estimated time trend may be used to adjust the concentration provided by an analysis which is carried out some period of time after actual sampling.

Suppose that a sample is analyzed at time  $t^*$  (in days). Let  $y(t^*)$  denote the resulting observation and let  $w(t^*) = \ln y(t^*)$ . If, for the particular contaminant under study, the estimated value of the parameter  $\alpha$  is  $a$ , the following relationship exists between  $w(t^*)$  and  $\hat{w}(0)$ , the (transformed) estimated observation at time  $t = 0$ , i.e., the time of sampling:



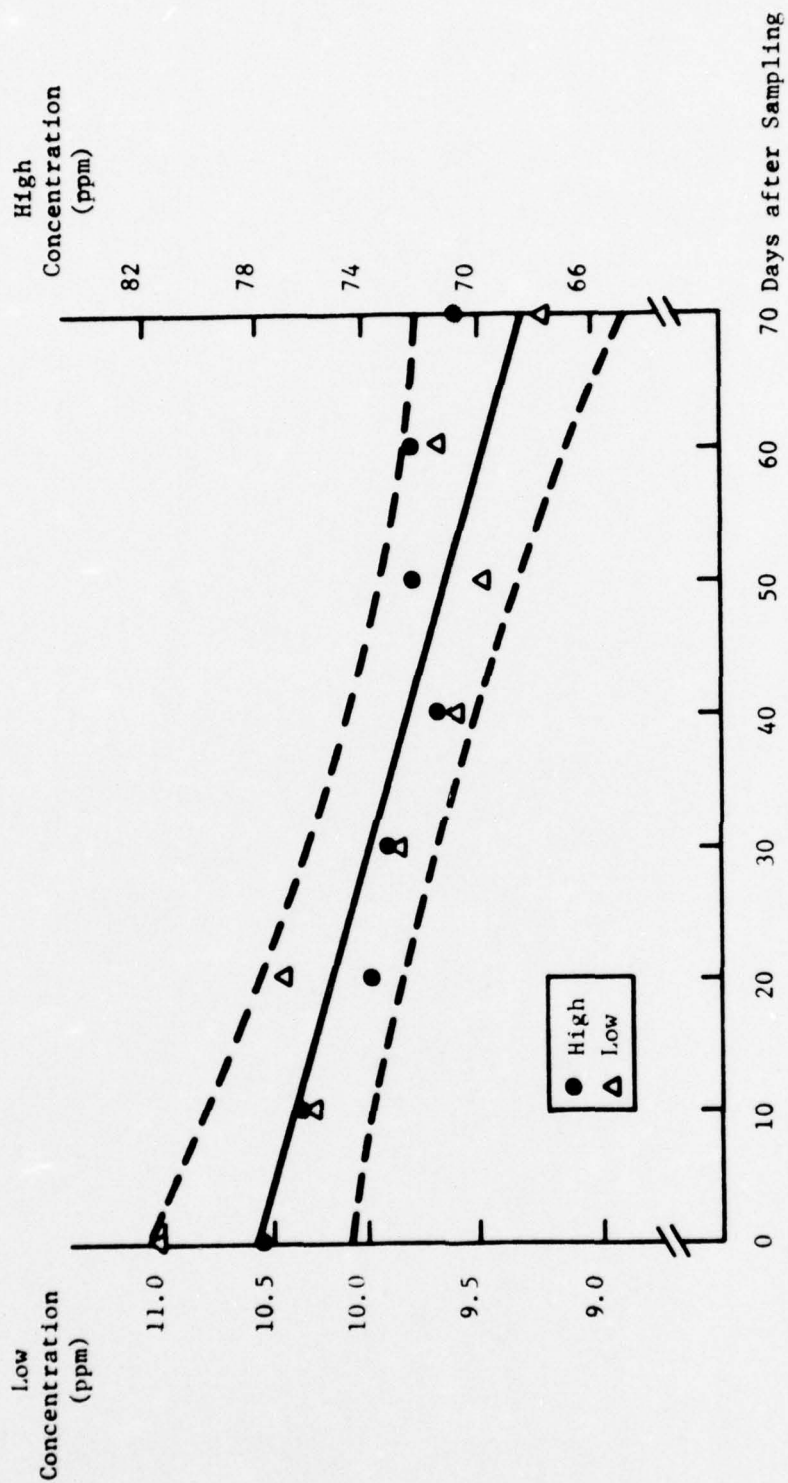


Figure 3: Time Trend for Contaminant #2 (Carbon Monoxide)  
with 95% Confidence Bounds

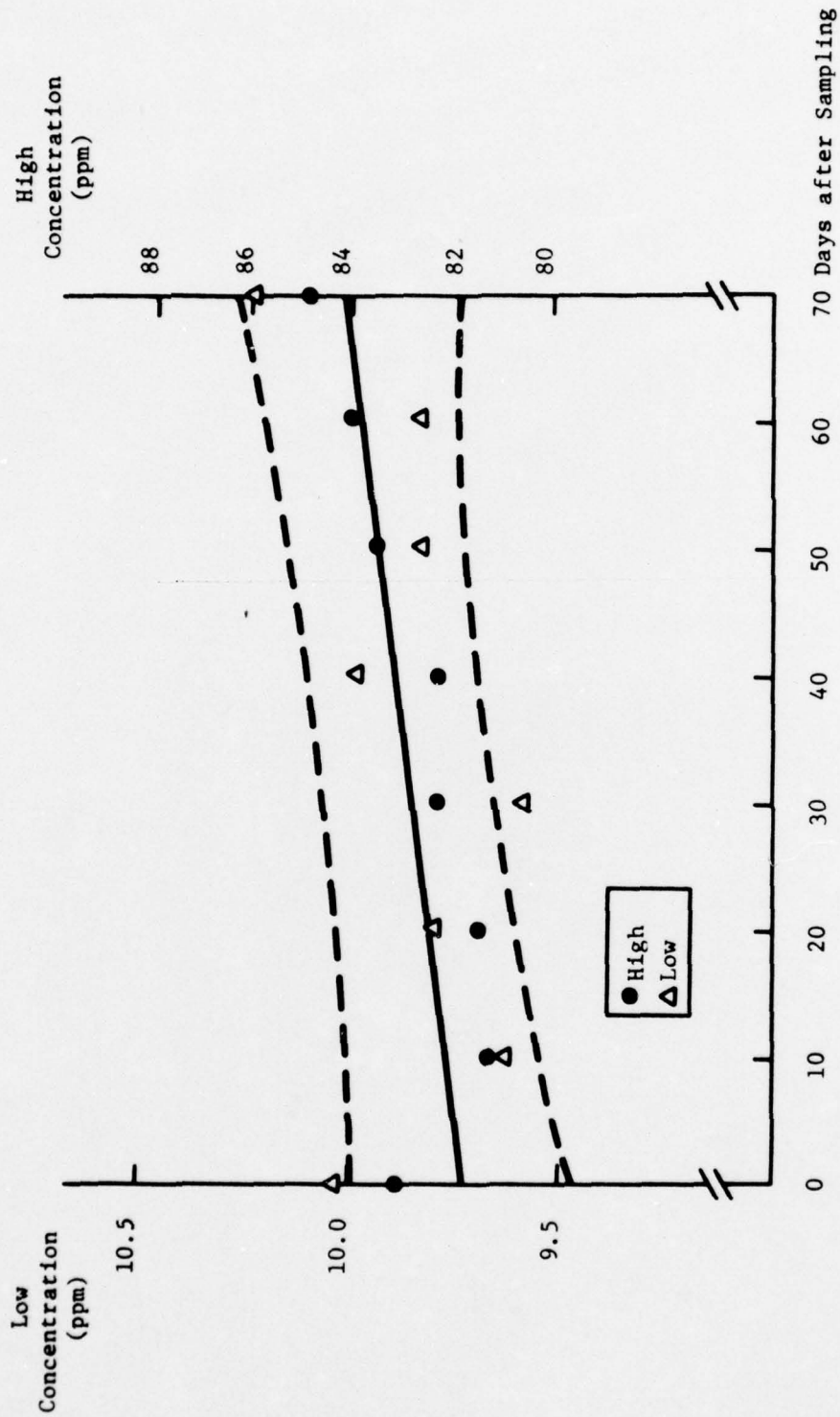


Figure 4: Time Trend for Contaminant #3 (R-12)  
with 95% Confidence Bounds

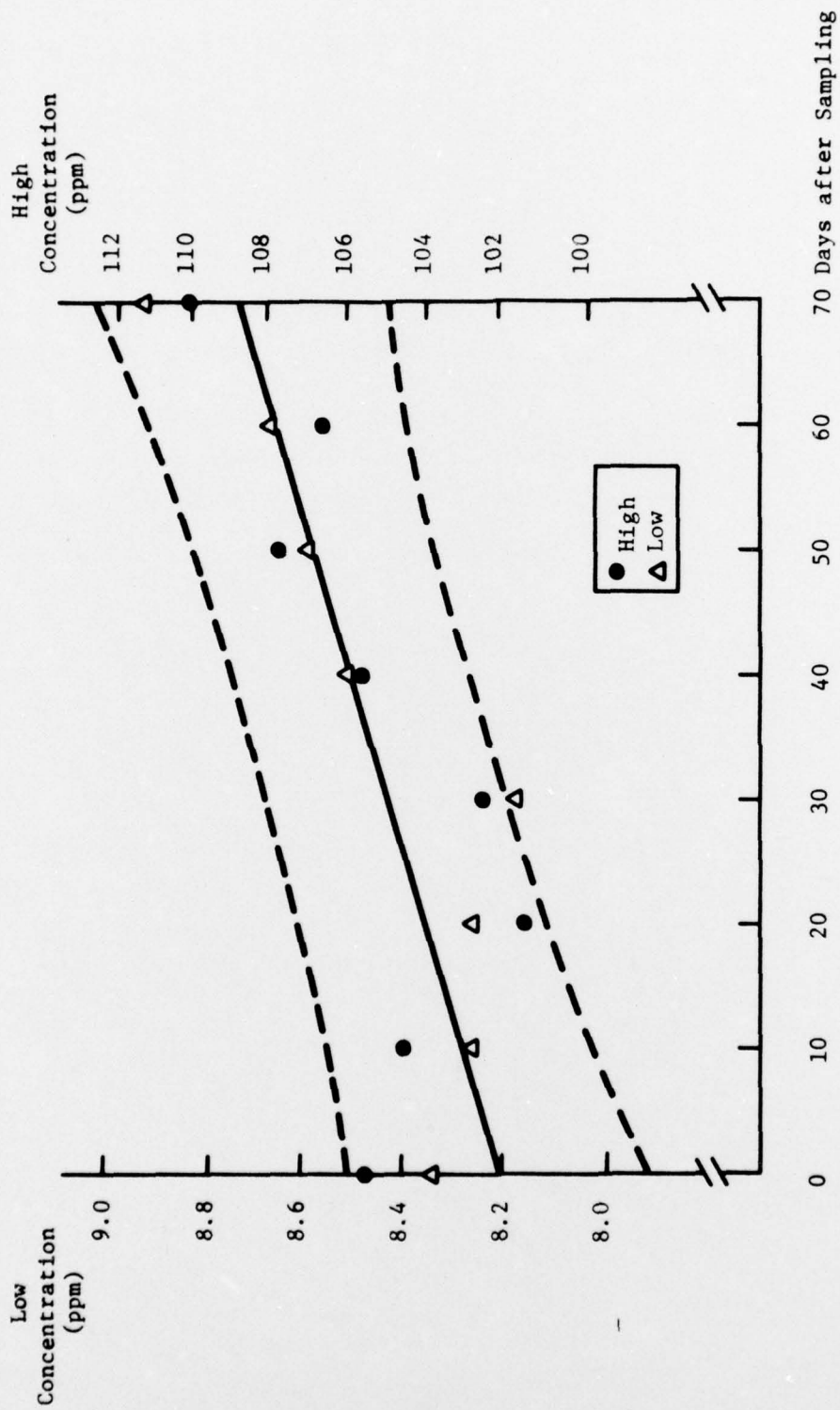


Figure 5: Time Trend for Contaminant #4 (R-114)  
with 95% Confidence Bounds



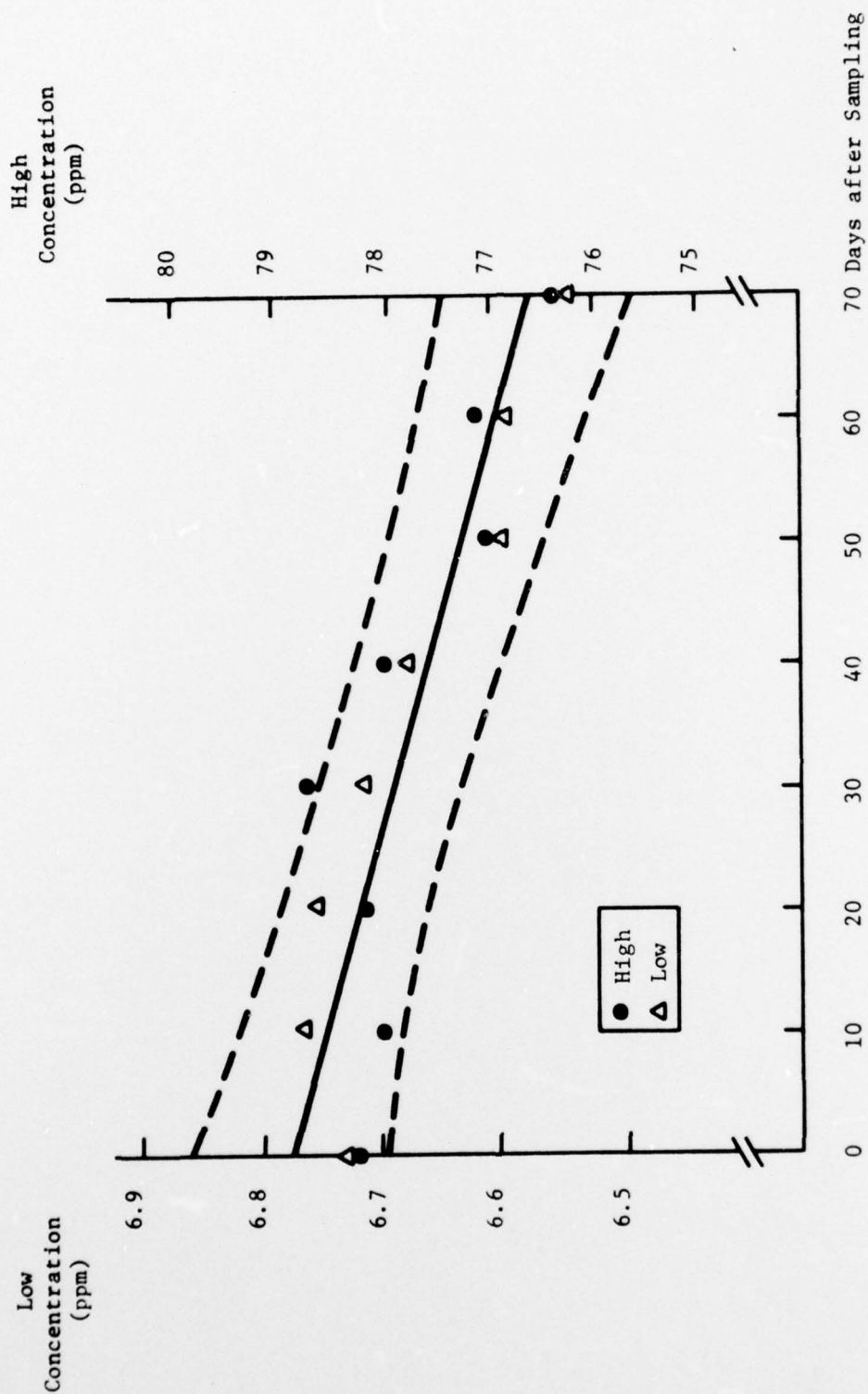


Figure 6: Time Trend for Contaminant #5 (Vinylidene Chloride)  
with 95% Confidence Bounds

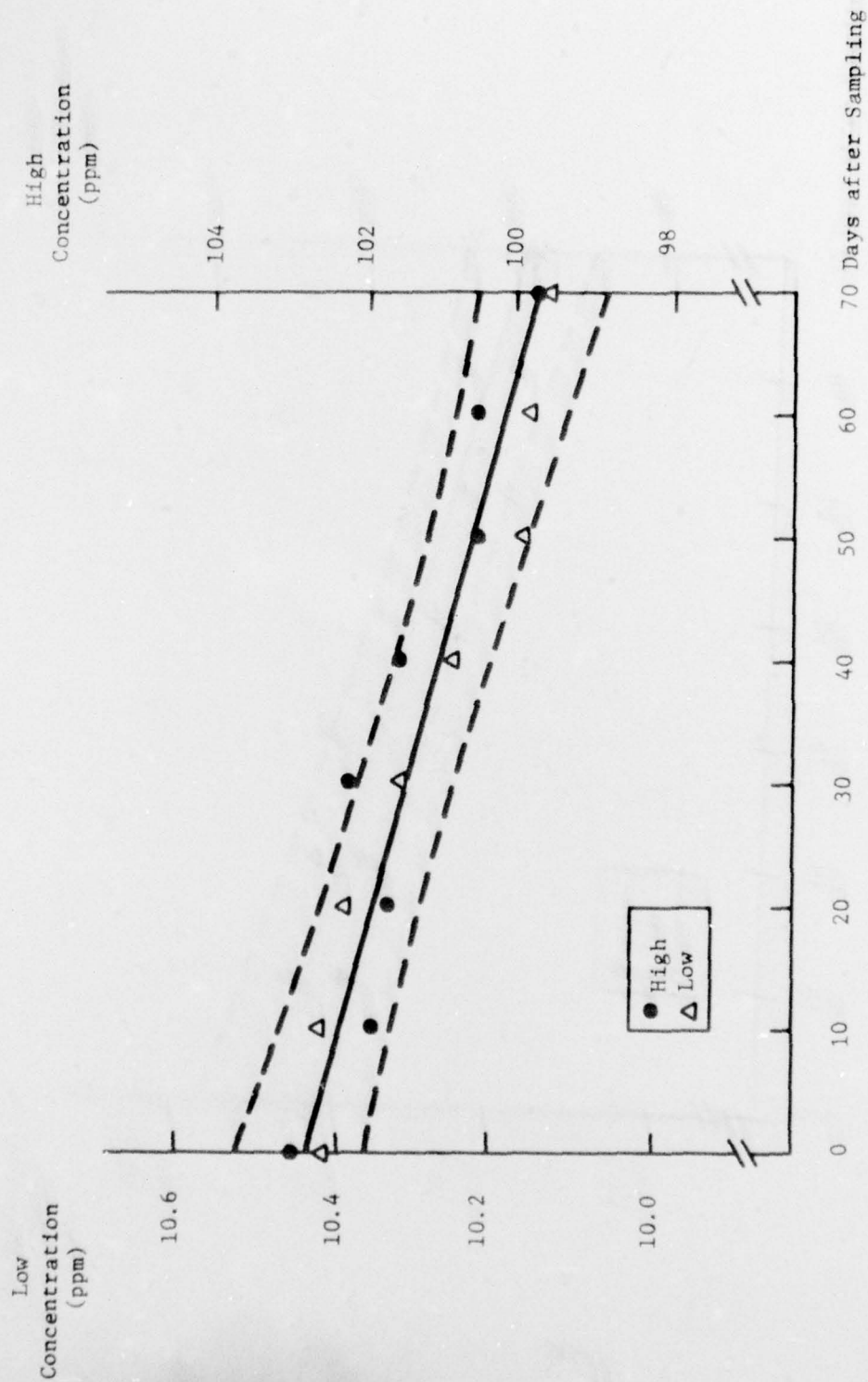


Figure 7: Time Trend for Contaminant #7 (Hexane)  
with 95% Confidence Bounds

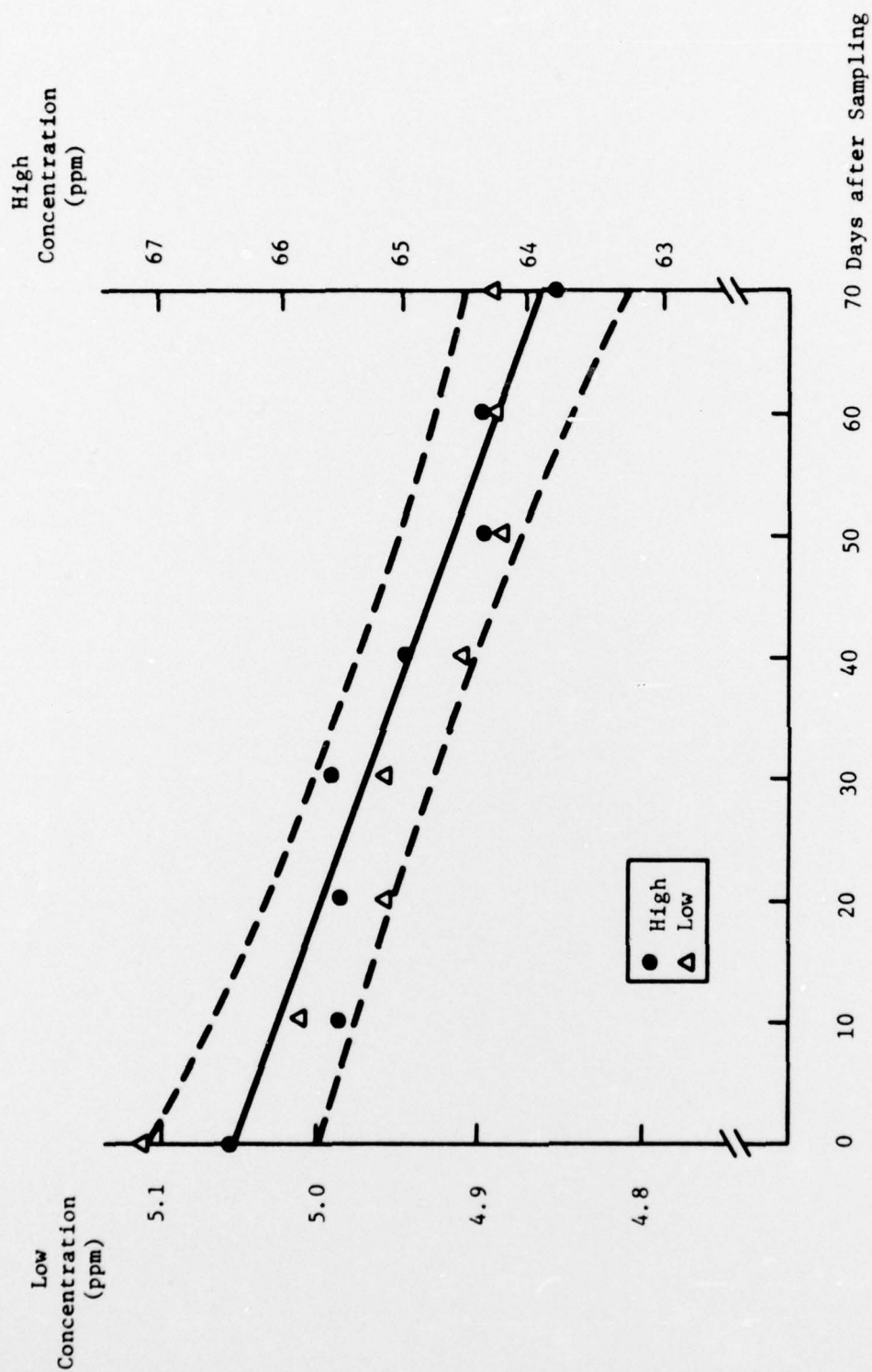


Figure 8: Time Trend for Contaminant #8 (Methyl Chloroform)  
with 95% Confidence Bounds

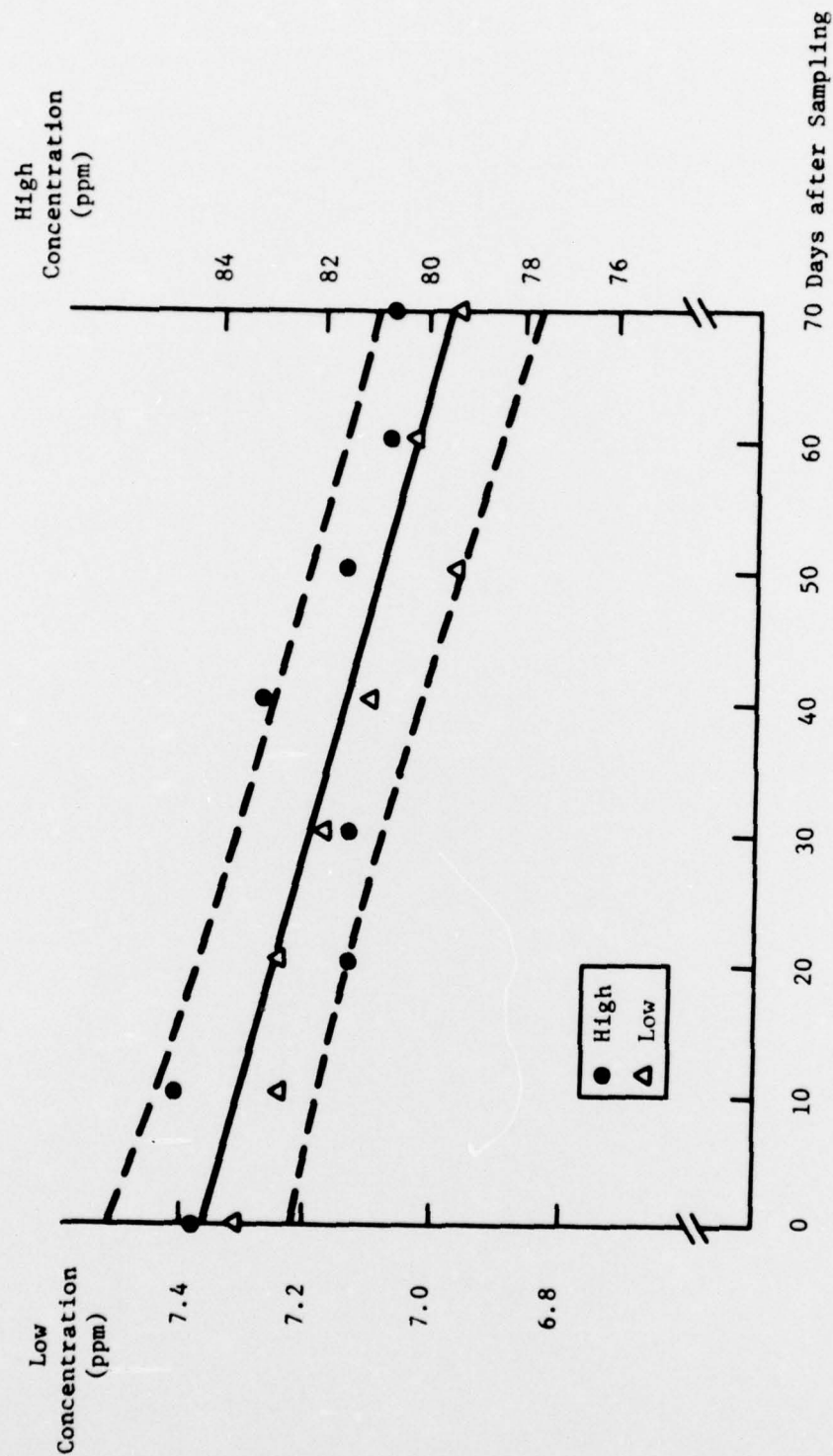


Figure 9: Time Trend for Contaminant #9 (Benzene)  
with 95% Confidence Bounds



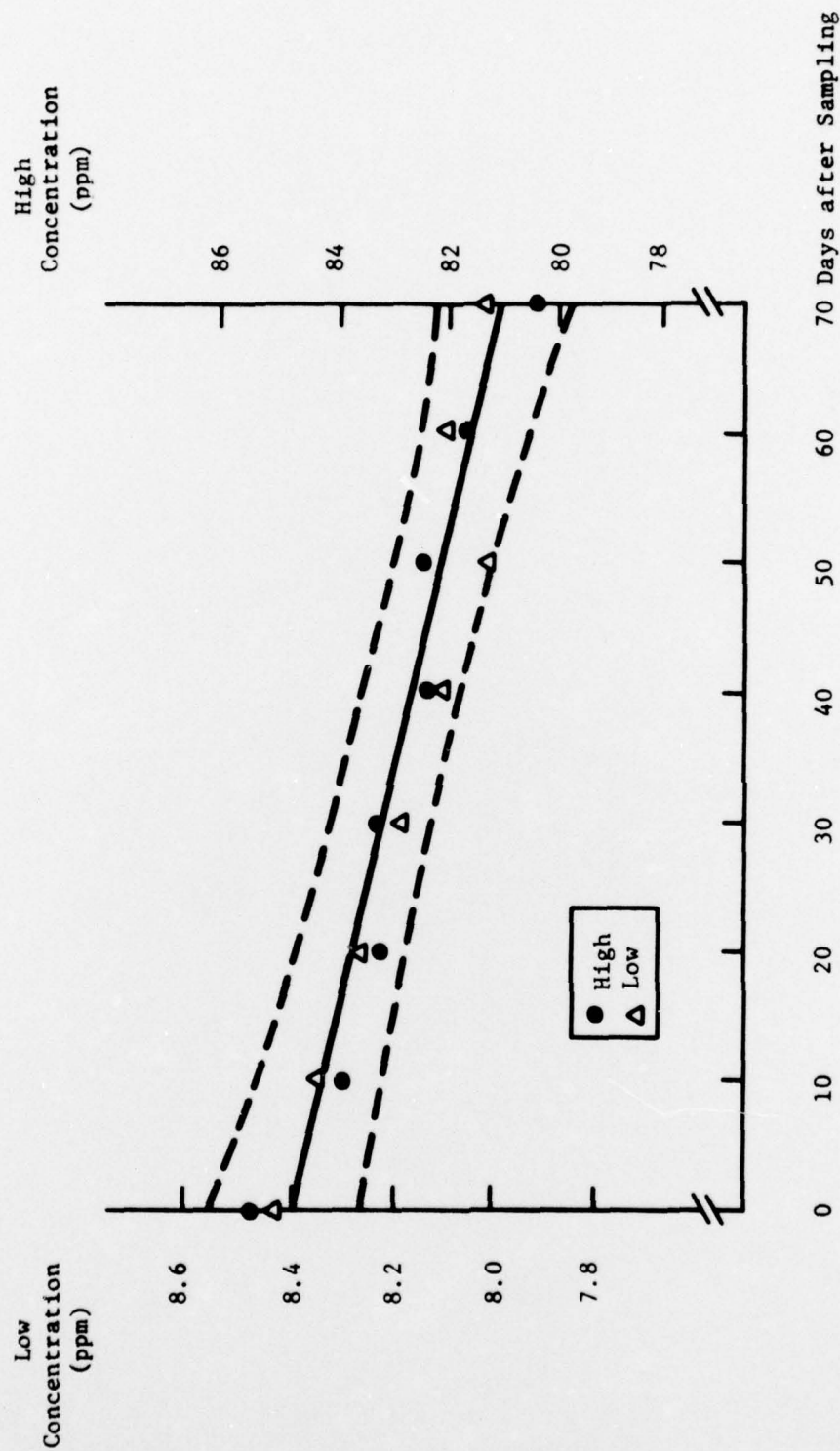


Figure 10: Time Trend for Contaminant #10 (Trichloroethylene)  
with 95% Confidence Bounds

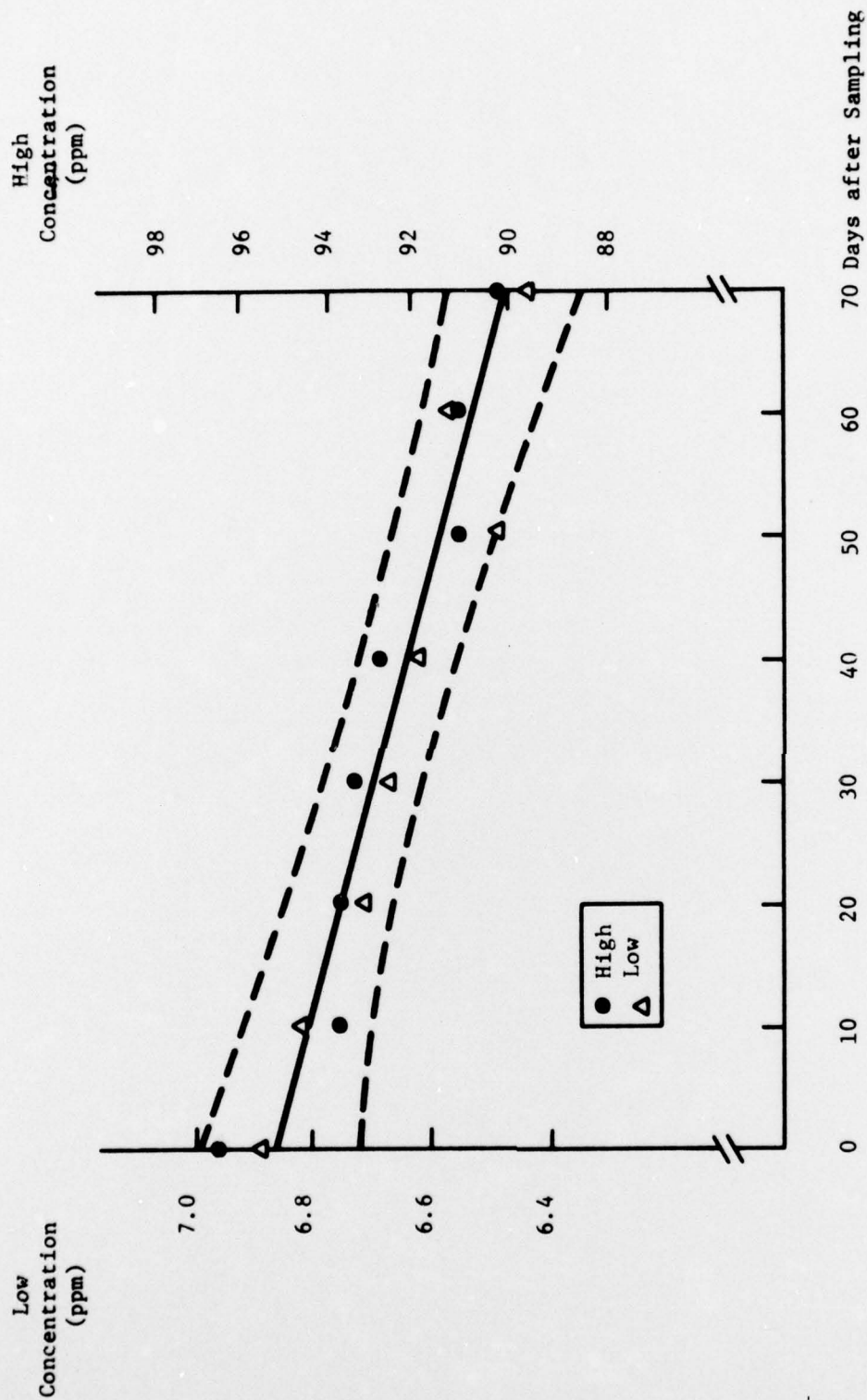


Figure 11: Time Trend for Contaminant #11 (Toluene)  
with 95% Confidence Bounds

$$E[\hat{w}(0)] = E[w(t^*) - at^*] .$$

Thus, as intuition indicates, an observation at time  $t^*$  would be adjusted by extrapolating backward along the regression line to time  $t = 0$ , and then untransforming the "logged" observations. That is, the adjusted value would be given by

$$\hat{y}(0) = \exp[\ln y(t^*) - at^*] .$$

For example, suppose an analysis done on contaminant #11 (toluene) seventy days after the sample was taken showed a concentration of 84.9 ppm. The best estimate of the actual concentration of toluene when the sample was taken would be:

$$\begin{aligned} \hat{y}(0) &= \exp[\ln (84.9) - (-.000825) (70)] \\ &= \exp[4.44147 + .05775] \\ &= \exp[4.49922] \\ &= 90.0 \text{ ppm} . \end{aligned}$$

Although this value (90.0 ppm) provides the best estimate of the actual concentration at time  $t = 0$ , it must be realized that there is a large degree of variability in this estimate. This variability results not from the variability in the regression line, which is minimal, but rather from the inherent variability in a single observation. This can be seen from the approximate 95% confidence interval for  $\hat{y}(0)$ .

It can be verified that the variance of  $\hat{w}(0)$  is given by

$$\begin{aligned} V[\hat{w}(0)] &= V[w(t^*) - at^*] \\ &= V[w(t^*)] + (t^*)^2 V(a) - 2 \text{ cov}[at^*, w(t^*)] \\ &= \sigma^2 [1 + (1 + [(t^*-35)^2 + (t^*)^2]/4200)/48] . \end{aligned}$$

From the experiment an estimate of the value of  $\sigma^2$  may be obtained from the plot x time (PT) mean square. For toluene, the estimate of  $\sigma^2$  is .0022, as can be seen from Table 15. Since this estimate has 28 degrees of freedom, the t-distribution may be used to construct a 95% confidence interval for  $\hat{w}(0)$ .

Skipping over the algebra, the resulting 95% confidence interval for the value of  $\hat{w}(0)$  in the example is

(4.40185, 4.59660)

Thus, an approximate 95% confidence interval for  $\hat{y}(0)$ , the actual concentration on the original scale, is

(81.6, 99.2)

This relatively wide confidence interval is, of course, reflecting the variability of a single sample. This interval may be made narrower, however, by obtaining more than a single sample in a given situation. For example, if the observed concentration of 84.9 ppm were based on the average of four samples, the 95% confidence interval for  $\hat{y}(0)$  would be reduced to (approximately):

(87.0, 94.4) .

#### C. VARIABILITY OF THE SAMPLING METHODS

As mentioned in the previous section, the analysis of variance indicated that no differences existed between the four sampling methods. This conclusion pertains to differences in means. The analysis of variance does not, however, indicate whether the variability of the observations



within any given sampling method is different than the variability of another sampling method.

Because this is of interest, tests for the quality of variance were conducted using the Hartley  $F_{\max}$  statistic which is a ratio of the largest to the smallest of the four sample variances. In general, no statistically significant differences in variability of the four sampling methods were noted. However, for some reason, during the initial analysis (i.e., at time  $t = 0$ ), measurement error tended to be smaller for sampling method #2 (vacuum only, single purge) than for the other sampling methods.

#### D. VARIATION DUE TO FLASKS

In the experiment, the variability of any observation is assumed to be composed of a number of sources. One of these sources is variability resulting from the fact that different flasks are used in the experiment and are assumed to be a random factor. In other words, the flasks are treated as a random sample from a population of flasks, rather than as flasks which are specifically identified. This treatment is realistic since, in general, the actual flasks used in sampling are obtained "off the shelf".

If, on the other hand, the flasks were treated as a fixed factor, where each individual flask had an unique identity, the variance of any observed observation would be reduced, because the singular effect of each flask could be accounted for.

In the current experiment, variation due to flasks appears in two components ( $F$  and  $F_T$  of equation (1) in section II.B.) Because of the consideration of only one flask for low concentrations, estimates of

these components were obtained only for high concentrations. For all contaminants other than #10 (trichloroethylene), at least one of these two components was statistically significant, indicating that using different flasks does add variability to the results.

In order to get some idea of how much random variability might be eliminated by regarding flasks as a fixed factor, the estimate  $V_F$  of the variability of an observation with flasks regarded as fixed was compared with the estimate  $V_R$  of the variability of an observation with flasks regarded as random. (For purposes here, the variation of interest is about the plot mean. Thus, variation due to plots was not included in the calculations.) A rough estimate of how much the variability may be decreased is given by the ratio  $(V_R - V_F)/V_R$  which ranges from a low of 6% for contaminant #2 (carbon monoxide) to a high of 77% for contaminant #8 (methyl chloroform).

#### IV. SUMMARY AND CONCLUSIONS

The two main topics addressed in this experiment were (1) whether differences existed in the performance of the four grab sampling methods used, and (2) whether the grab samples have a shelf life, i.e., whether a delay in analyzing the samples affects the results. The statistical analysis of the experimental data indicates that each of the four sampling methods performs equally well. Thus, the choice of one of these grab sampling methods should be made on some other basis, such as the time and costs associated with preparation and actual sampling.

Although the experiment did not reveal any differences in grab sampling methods, it did indicate that the samples do have a shelf life. For all contaminants except the refrigerants and methane, a statistically significant exponential time decay is present in the shelf life of a sample. Although decay parameter values are not the same for each contaminant, the parameter value is independent of the concentration. For two of the refrigerants (R-12 and R-114), a statistically significant exponential time growth was found in the shelf life, again independent of concentration. No statistically significant time effect existed for the remaining refrigerant (R-113) or for methane. For those contaminants which exhibited a statistically significant time effect in shelf life, a better estimate of the actual contaminant concentration at the time of sampling may be obtained by adjusting the observed value as indicated in the previous section.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 106-4	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A STATISTICAL STUDY OF GRAB SAMPLING IN A CLOSED ENVIRONMENT		5. TYPE OF REPORT & PERIOD COVERED Technical Report
6. AUTHOR(s) Dennis E. Smith and Terry L. King		7. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Desmatics, Inc. P. O. Box 863 State College, Pa. 16801		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-1054
11. CONTROLLING OFFICE NAME AND ADDRESS Statistics and Probability Program (Code 436) Office of Naval Research Arlington, Va. 22217		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-334
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 52p.		12. REPORT DATE February 1977
		13. NUMBER OF PAGES 47
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Distribution of this report is unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) 14 TR-106-4		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Atmospheric contaminants Atmospheric sampling Grab sampling Statistical analysis		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Maintaining the quality of an enclosed environment (e.g., aboard a sub- marine) requires monitoring of the environment by means of sampling procedures. One general type of sampling procedure is "grab sampling", in which atmospheric samples are obtained, essentially instantaneously, in evacuated flasks. This report describes an experiment to determine (1) whether differences exist in the performance of four grab sampling methods, and (2) whether a delay in analyzing the samples affects the → next page		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

391156

VB




UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

cont.

→ results (i.e., whether the samples have a shelf life).

Eleven contaminants, each at a high and a low concentration level, were used in the study. Grab samples of these contaminants were analyzed at ten day intervals from the time sampling occurred until seventy days had elapsed. Although the experiment did not reveal differences in grab sampling methods, it did indicate that, for most of the contaminants, the samples do have a shelf life.



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)